Editorial

Video Analysis, Abstraction, and Retrieval: Techniques and Applications

Jungong Han,1 Ling Shao,2 Peter H. N. de With,1,3 and Ling Guan4

1 Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands
2 The University of Sheffield, Sheffield S1 3JD, UK
3 CycloMedia Technology, 418 0BB Waardenburg, The Netherlands
4 Ryerson University, Toronto, ON, Canada M5B 1Z2

Correspondence should be addressed to Jungong Han, jg.han@tue.nl

Received 9 May 2010; Accepted 9 May 2010

Copyright © 2010 Jungong Han et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The proliferation of TV programs and personal DV cameras has led to an explosion of digital video content, which enriches the personal entertainment of users. However, the rapidly increasing availability of video data has not yet been accompanied by an increase in its accessibility. This is due to the fact that video data are naturally different to traditional forms of data, which can be easily accessed and searched using textual queries. Therefore, the problem of efficiently organizing video, such as TV news and sports, into more compact forms and extracting semantically meaningful information becomes more and more important. In the past ten years, video analysis and retrieval techniques have received significant attention from both industry and academia. The research has gradually converged to three fundamental areas, namely, video analysis, video abstraction, and video retrieval. Video analysis is utilized to extract both the general and the domain-specific visual features, such as colour, texture, shape, human faces, and human motion. Video abstraction generates a representation of visual information, similar to the extraction of keywords or summaries in text-document processing. Basically, video abstraction is associated with key-frame detection, shot clustering, and the extraction of domain knowledge in a video source. The content attributes found in video analysis and abstraction processes are often referred to as metadata. Video retrieval based on the extracted metadata is a fast and interactive tool that allows users to query, search, and browse large video databases. Although a lot of efforts have been devoted into these three areas, both the computational cost and the accuracy of the existing systems are still far from satisfactory. For example, a typical problem is how to optimally use unlabeled data in case that the number of available training samples is very small. Another problem is that in a general setting, the low-level features do not have a direct link to high-level concepts. This raises the question how this semantic gap can be bridged.

This special issue highlights the most recent advances and promising results of the research community working in video analysis, abstraction, and retrieval. After two rounds of careful reviews, nine papers were selected for publication in this special issue. We group the selected papers into three thematic sections corresponding to three research areas mentioned above, though we notice that some papers cover two or even three areas.

The first section deals with feature extraction issues in video analysis. The first paper entitled “ipProjector: designs and techniques for geometry-based interactive applications using a portable” presents an interactive projection system for virtual studio setup using a single self-contained and portable projection device. The projection allows special effects of a virtual studio to be seen by live audiences in real time. The techniques discussed in this paper, such as colour wheel analysis and motion-based camera calibration, are highly related to and widely used in the video analysis area. The second paper, entitled “Flexible human behaviour analysis framework for video surveillance applications,” investigates the human motion based on a two-camera setup, which is a key clue for analyzing the surveillance video. The main contribution is the effective combination of trajectory estimation and human-body modelling, facilitating
the semantic analysis of human activities in video sequences. Moreover, an automatic camera-calibration technique is employed to establish the correspondence between the two video channels. By doing so, the system can make decision based on fusing the information from both cameras, thereby resulting in a robust detection.

The second section discusses the video structuring, video segmenting, and shot detection, which are all hot topics in the video abstraction area. This section includes four papers. The first paper entitled “Statistical skimming of feature films” presents a statistical framework based on Hidden Markov Models (HMMs) for skimming feature films. This work combines the information derived from the story structure with the characterization of the shots in terms of salient features. The structure of the video is captured by HMMs, which model semantic scenes and produce the shot sequence of the final skim. The second paper entitled “An optimized dynamic scene change detection algorithm for H.264/AVC encoded video sequence” concentrates on scene change detection for compressed video sequences. The scene detector employs a dynamic threshold that adaptively tracks different features of the video sequence, thereby increasing the accuracy in correctly locating true scene changes. The work has been successfully applied within an error-concealment framework for H.264 decoding. The third paper entitled “Automatic TV broadcast structuring” proposes a fully automatic system to detect the start and end times of each program in TV broadcasts. The algorithm is based on the detection of repeated sequences, in order to extract long useful programs, such as movies, news, TV series, and TV shows. The last paper in this section entitled “Unsupervised segmentation methods for TV contents” also deals with the shot boundary detection problem. This paper analyzes both the audio and the video signal of a sequence, rather than relying on the video signal only. The basic system is built upon the hypothesis that it is possible to segment any audiovisual document into homogeneous segments at the adequate scale. The system has been evaluated for two different applications: TV program boundary detection and speaker diarization. For both cases, the system achieves high accuracy.

The third section addresses the video indexing, browsing, and retrieval. The first paper entitled “A video browsing tool for content management in postproduction” implements an interactive video browsing tool for supporting content management and selection in postproduction. Many visual features, like camera motion, visual activity, face occurrence, global colour similarity, and repeated takes, are extracted and used in this system. The second paper entitled “Personalized sports video customization using content and context analysis” focuses on sports videos and addresses three research issues: semantic video annotation, personalized video retrieval, and system adaptation. The system is designed for users to watch refined video segments containing their favourite semantics instead of lengthy sports matches. Moreover, both subjective content preference and objective environment constraints are well balanced so that the optimal visual experience can be brought to the particular viewer. The last paper “Multimodal indexing of multilingual news video” deals with the analysis of multilingual news telecasts in India. The basic approach is to index the news stories with relevant keywords discovered in speech and in form of “ticker text” on the visuals. They also create a multilingual keywords-list in English and Indian languages, to enable keyword spotting in different TV channels, both in spoken and visual forms. The evaluation shows that restricting the keyword list to a manageable size results in drastic improvement in indexing performance.

We would like to thank all the authors for sharing with us their nice and innovative work. We also express our sincere gratitude to all the reviewers for their timely and insightful comments on selecting papers. Finally, we are particularly grateful to the Editor-in-Chief, Dr. Fa-long Luo, for his invitation and guidance throughout the entire process. Without his support, we could not make this special issue possible.

Jungong Han
Ling Shao
Peter H. N. de With
Ling Guan